

original research report

Intersite validations of the pixel-wise method for liver R2* analysis in transfusion-dependent thalassemia patients: a more accessible and affordable diagnostic technology

Pairash Saiviroonporn,^a Vip Viprakasit,^b Kleebasabai Sanpakit,^b John C. Wood,^c Rungroj Kittayaphong^d

From the ^aDepartment of Radiology, ^bHaematology/Oncology Division, Department of Pediatrics and Thalassemia Center, Bangkok Thailand, ^cDivisions of Cardiology, Department of Pediatrics, Children's Hospital Los Angeles, Keck School of Medicine, University of Southern California, Los Angeles, California, United States and ^dDivision of Cardiology, Department of Internal Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand

Correspondence: Vip Viprakasit, MD · Division of Hematology/Oncology, Department of Pediatrics and Thalassemia Center, Faculty of Medicine Siriraj Hospital, Mahidol University, 2 Prannok Road, Bangkoknoi, Bangkok 10700, Thailand · T/F: +66-4122113 · sivvp@mahidol.ac.th

Hematol Oncol Stem Cell Ther 2012; 91-95

DOI:10.5144/1658-3876.2012.91

BACKGROUND AND OBJECTIVES: MRI-R2* has been accepted as a clinical tool for monitoring iron overload in thalassemia patients, especially for monitoring liver iron concentration (LIC). The most optimal and practical method of analysis however, is still open to further investigations. Our objective was to investigate intra- and intersite observer variability of the pixel-wise method for liver R2* analysis in thalassemia patients using a mono-exponential with a constant offset model.

PATIENTS AND METHODS: We performed 88 liver R2* measurements on 72 thalassemia major patients. A single breath-hold multi-echo gradient-echo sequence was acquired and analyzed at both the reference (REF) and local (LOC) sites. The analysis defined the region of interest in the whole liver parenchyma, excluding the great vessels, and were reported as median values.

RESULTS: The R2* values from the REF and LOC were statistically comparable for all comparisons. The intrasite and intersite observer variation were 0.75% (less than 0.9%) and 2.5%, respectively, both of which are comparable to previous reports, but substantially lower than conventional region-based approaches.

CONCLUSION: The low variation of the R2* also yielded excellent variation in the tabulated hepatic iron content. However, caution is required when comparing the results to different implementation methods and appropriate evaluation and validation of methodology for any new scan site is essential before its clinical use.

The R2* measurement is recently gaining more acceptance as a clinical tool for monitoring iron concentration in thalassemia patients.¹⁻³ The measurement can be conveniently performed on various organs, such as the heart, liver or pancreas.⁴⁻⁷ Liver iron provides considerable information about total body iron stores and risk of iron-mediated complications.⁸⁻¹¹ Furthermore, R2* measurements also have the benefit of enabling the monitoring of the efficacy of the chelator agents whose efficacies on various organs are different.¹²⁻¹⁵

The measurement can be divided into two parts: magnetic resonance (MR) acquisition and post-processing analysis. Currently, the MR acquisition part is, well established in the clinical arena due to its robust-

ness and can be scanned within a single breath-hold time.^{3,16} On the other hand, the analysis part involves various implementation methods. The differences may vary from calculation approaches, fitting models, or region of interest (ROI) selections.³ In general, there are two main calculation approaches: the region-based (RB) and pixel-wise (PW) methods,^{3,4,17} which can be analyzed on the single breath-hold images. The RB method begins with the averaging of each echo signal using a defined ROI and then fitting it to the exponential models. The PW method, on the other hand, fits the decay signal point by point over the whole liver and then averages the R2* values from the defined ROI. The RB method is a fast and robust technique, which is less

susceptible to respiratory motion, but it cannot provide the distribution pattern of R2* values in the liver. The PW method can overcome this limitation, but is more computationally intensive. In terms of fitting methodologies, there are three accepted models for fitting the decay curves: the single-exponential (S-EXP), mono-exponential with constant offset (C-EXP), and bi-exponential (BI-EXP) models.^{3,4,17,18} The S-EXP model can be implemented with or without truncating the plateau part of the signal from the fitting (truncation model). Finally, the ROI can be defined into two separate strategies as from partial or whole liver parenchyma.¹⁹ Previously, Positano et al²⁰ investigated all such the differences on simulated data as well as on thalassemia patients and found that using median R2* calculated by the PW approach with the C-EXP model on the whole liver ROIs yielded the lowest intra- and interobserver variations. However, there was some limitation since this study is performed only on-site with custom automatic segmentation software for ROI generation.

Determination of liver iron concentration (LIC) has long been recognized as a standard monitoring in transfusion-dependent thalassaemia (TDT) patients who receive regular blood transfusions.²¹ Until recently it has been shown that the MRI approach for LIC measurement is also critically important in nontransfusion-dependent thalassaemia (NTDT) patients who show a weaker correlation between serum ferritin (SF) and LIC. Therefore, SF might not be able to predict the level of liver siderosis and guides iron chelation practice as it does in TDT.¹¹ Moreover, using a non-invasive approach by MRI measurement can reduce undesirable complications and uncertainty related to standard liver biopsy. Although such an MRI measurement has been available by means of a standard service provider such as Ferriscan (Resonance Health Limited and Resonance Health Analysis Services Pty Limited, Claremont, Australia) in many developed countries, this essential monitoring remains limited to only a few thalassaemia patients, especially in the Far East where the cost of imaging and data interpretation is still unaffordable by most patients.²² To provide accessibility to more affordable new technology to under-funded thalassaemia patients in a setting where resources are limited such as Thailand, we have established the North-South collaboration between Mahidol University, Bangkok and the University of Southern California, Los Angeles, California, United States. Our local software designed to determine cardiac T2* has been established and validated as reported previously.³ In this report, we describe our further

evaluation of the software on the determination of LIC using MRI-R2*. We investigated intersite variation of the reference site (Children's Hospital Los Angeles, California, United States) and our institute (Siriraj Hospital, Bangkok, Thailand) using manually defined ROI and independently generated analysis tools.

PATIENTS AND METHODS

Eighty-eight liver R2* measurements were performed on 72 transfusion-dependent thalassemia (TDT) patients (32 males and 40 females, mean [SD] age 18.0 [6.9] years) who received regular transfusion and iron chelation therapy since ≥ 10 years old and had serum ferritin levels more than 1000 ng/mL. Patients with a contraindication to MR (such as pacemakers and claustrophobia) and an inability to comply with the instruction were excluded from the study. Approval of the study was obtained from the Institutional Review Board of Faculty of Medicine, Siriraj Hospital, Mahidol University.

The study was conducted on a 1.5T Philips Achieva XR Quasar Dual Gradient system (Philips Medical Systems, the Netherlands) using a five-element cardiac phased-array coil. Each patient was scanned at a mid-hepatic slice using a multi-echo fast gradient-recalled-echo sequence acquired within a single breath-hold time.⁵ Imaging parameters were a repetition time (TR) of 80 ms, 20 echo times (1.07-16.30 ms with 0.80 ms increments), a flip angle of 20 degrees, a slice thickness of 10 mm, a matrix of 128 \times 256 pixels, and a field of view (FOV) of 40 cm, which yielded a voxel size of 3.1 \times 1.6 \times 10 mm³. In this study, hence, the maximum practical measurable R2* due to the limitation of MR acquisition was 1308 Hz, which is a reciprocal of the minimum TE employed, divided by 1.4.^{3,20}

The acquired images were analyzed locally (LOC) and also transferred in DICOM format to process at the reference (REF) sites. All R2* analyses were performed independently at both sites using individually-developed custom-written software implemented in MATLAB software (The MathWork, Natick, Massachusetts, United States).³ The PW method using the C-EXP model was employed for such calculation at both sites. The raw and R2* parametric images were then employed to manually define a large ROI of the whole liver parenchyma excluding the major blood vessels.^{3,17} Two MR technologists independently performed the analysis at the local site (intrasite observer) and data were also compared to the result from the reference site (intersite observer). The R2* results were reported as the median of the data inside the defined ROI and the hepatic iron content (HIC) was

tabulated from the linear relationship with the R2* according to the expression of (3):

$$\text{HIC (mg/g dw)} = 0.0254 \times \text{R2* (Hz)} + 0.202$$

According to the maximum practical measurable R2* in this study, the maximum measurable HIC in this study, hence, was about 33.40 mg/g dry weight.

All statistical analyses were evaluated using the MATLAB Statistics Toolbox. A Bland-Altman plot was used to analyze the agreement between the two different R2* data sets. The coefficient of variation (CV) was used as the quantitative analysis of the closeness of the agreement. The CV was calculated from the standard deviation of the difference between the two data sets, then divided by their mean and presented as a percentage. The paired *t* test was selected to evaluate the difference between the data. A *P* value less than .05 was considered to be significant.

RESULTS

An example of liver images with corresponding tabulated HIC map overlaid with defined ROIs (the blue line) from the whole liver parenchyma are shown in **Figure 1**. These data were obtained from a thalassemia patient who had severe iron concentration in the liver.

In this study, the mean R2* values from the reference and local sites were 695 ± 463 and 700 ± 463 Hz, respectively (*P*=not significant). The differences between the intra- and interobservers at the local site were also not significant. Sixteen patients were excluded from the validation because their liver R2* was out of the measurable range ($\text{R2*} > 1308$ Hz).

The Bland-Altman plots of the R2* and HIC values from intra- and intersite observers of the measurable data range are shown in **Figure 2**. The corresponding bias, 95% CI, and the CV of these variations are presented in **Table 1**. The intrasite variation in the measurable range had a small bias (0.39 Hz) and CV (0.83%). The intersite variation was also remarkably low with R2* bias less than 5 Hz and CV of 2.5% as shown in Figure 2 and in the last row of **Table 1**. Thus, the variation of the tabulated HIC from the intersite variability in the study had a bias lower than 0.12 mg/g dry weight and CV of 2.2%. The linear correlation of R2* between the two sites was in very good agreement with a coefficient of determination of 0.998. However, if all patient data were included for the validation, the intersite variation would have a bias of about 8 Hz and 4.5% of CV with a coefficient of determination of 0.986. In summary, in this study, the intrasite CV of the R2* data in the measurable range was less than 0.9%, while it was only about 2.5% for the intersite with a low bias for all comparisons (less

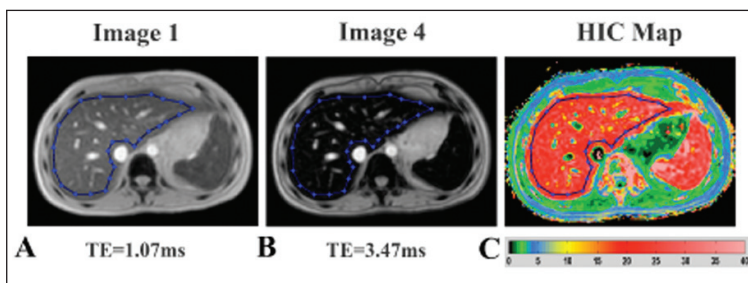


Figure 1. Example liver images from the first (A) and fourth (B) echoes and corresponding hepatic iron content (HIC) map (C) overlaid with ROIs (the blue line) obtained from a thalassemia patient. The color-map scale bar for hepatic iron content (HIC) map is displayed at the bottom in mg/g dry weight.

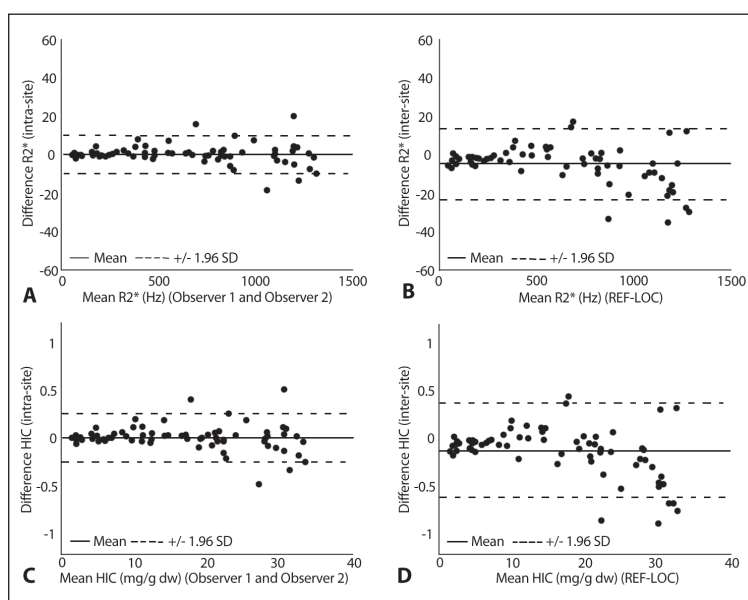


Figure 2. Bland-Altman plots of intra- and intersite reproducibility of the R2* and tabulated hepatic iron content (HIC) (the top and bottom row, respectively). The left column presents the intrasite observer variation while the intersite observer is on the right.

than 1 and 5 Hz for intra- and intersite variation, respectively). Such a low variation of the R2* yielded also an excellent variation in the tabulated HIC, which had a similar low bias and CVs.

DISCUSSION

At present, the MRI measurements of R2 and R2* have essentially replaced conventional liver biopsy for hepatic iron quantification in thalassemia patients, due to its accuracy and ease of use. Suitable R2* images can be obtained using a single breath hold multi-echo gradient-echo sequence that has become standard on all three major MRI platforms. Some disagreement still exists regarding the optimal analysis methods,³ although PW assessment with constant offset com-

Table 1. Bias, 95% CI, and coefficient of variation of intra- and intersite observers of the R2* and tabulated hepatic iron concentration calculated from the measurable data range while coefficient of variations from all measurable data are presented in the parentheses.

	R2*	Hepatic iron concentration		
	Bias [95% CI] (Hz)	Coefficient of variation (%)	Bias [95% CI] (mg/g dw)	Coefficient of variation (%)
Intrasite	0.39 [-9.64 10.44]	0.83 (1.31)	0.01 [-0.24 0.27]	0.80 (1.29)
Intersite	-4.72 [-23.38 13.93]	2.45 (4.46)	-0.12 [-0.60 0.35]	2.20 (3.58)

pensation has the largest validation against biopsy and against computer simulations.^{20,23} The analysis in this study, hence, employed such a method to further investigate the intra- and intersite variation and found that the method also provided comparable bias and variations as described in Positano's report,²⁰ even when using only a manually defined ROI instead of automatic liver segmentation. Intrasite variation was about 0.75%, which is comparable to the previously reported 0.85%, but substantially lower than 3.7% or 4.4% from the conventional RB approach.^{20,24} Intersite observer variability was about 2.5%, which is comparable to the same method reported by Positano et al at 2.9%, but lower than the 5.6% from the RB approach employed in clinical practice.^{20,24} The variability in this study arises from differences in tracing liver boundaries and differences in software implementation of the fitting routines; given the small overall error, these differences must be insignificant. While automatic segmentation may facilitate the speed of image interpretation, reproducibility was similar to manual boundaries.^{20,23}

Although the observer variation of the liver R2* analysis from the conventional RB approach was already in acceptable ranges, a two-fold reduction in variability may be important at high R2* values, for assessing interval change, and for clinical trial design; a two-fold lowering of variability translates to a four-fold decrease in patient accrual targets. However, caution must be used when comparing across studies. The lower intersite variation found in the present work might be due to high liver R2* values (723±480 Hz) which may have lower variability on a percentage basis than measurements on less heavily loaded patients. Prior reports have also reported T2* variability, which will tend to emphasize errors at lower iron concentrations. We also excluded patients whose R2* were greater than the measurable range (R2* >1308 Hz), which reduced the variation from 4.5% to 2.5%. Such high variations due to the out-of-range data have been reported in previous studies as well.^{3,20} The exclusion of such data, nevertheless, does not effect the clinical data range because it is more than two

times the cut-off point (R2*=588 Hz or HIC=15 mg/g dry weight) of what is considered to be severe iron overload in the liver.²⁵

Even though the study demonstrated a low R2* analysis, caution is still required when comparing the results to different implementation methods. Various factors may affect such comparisons. First, if we compared the PW to RB approach, the result will be comparable for the clinical range, but will be systematically underestimated by the RB method in the normal range (R2* <50 Hz).²⁰ Second, in our study we employed the C-EXP model so when comparing our result to another study using the BI-EXP model, the result is still fully comparable, but our R2* result will be underestimated by about 333 Hz over the whole R2* range when compared to the S-EXP model.²⁰ Finally, when comparing our result from the whole liver ROI to a partial one, as in the conventional method, the variation will be higher due to the sampling error on the latter method, which is due to the heterogeneity of live iron deposition.^{19,20}

The study may have some limitations. There are few patients who have liver iron overload in the borderline or normal ranges so the observer variation reported in this study relied heavily on patients with iron overload. More patients recruited in these borderline iron overload and normal LIC regions will help to better represent observer variation from the whole ranges. Furthermore, a basic semi-automatic liver segmentation should also help to further optimize the operator interactions and furthermore make the result more consistent for follow-up studies.

In conclusion, the intrasite CV of the liver R2* data in the measurable range was less than 0.9% in our study, while only about 2.5% for the intersite CV with a low bias for all comparisons. Such a low variation of the R2* yielded excellent variation in the tabulated HIC. Our "poor man" approach described in this study can reliably provide an accurate alternative for liver iron determination for clinical application to monitor and guide management in patients with iron overload who require iron chelation therapy.

Acknowledgement

The authors would like to thank all thalassemia patients and their parents who participated in this study. We thank Prajak Thanapiboonpol and Supaporn Nakyen for their technical assistance; Novartis Co., Thailand for its supports for MATLAB program; and Philips Medical Systems for their technical assistance.

Conflict of interest and funding

No conflict of interest. This work was supported by a grant from the Siriraj Grant for Research Development (R015233005) to R.K. and P. S. V.V. was supported by

Thailand Research Fund, BIOTEC and the National Research University Grant through Thalassemia Center, Mahidol University, Thailand.

Author contributions

PS was the principal investigator and takes primary responsibility for the paper. PS, VV, JCW, and RK conceived the study design and were involved in interpretation of data; KS supplied patient material; PS performed statistical analyses and wrote the manuscript and VV was responsible for the final draft of the manuscript. All authors read and approved the final version of the manuscript.

REFERENCES

- Angelucci E, Barosi G, Camaschella C, Cappellini MD, Cazzola M, Galanello R, et al. Italian Society of Hematology practice guidelines for the management of iron overload in thalassemia major and related disorders. *Haematologica*. 2008 May;93(5):741-52.
- Fischer R, Harmatz PR. Non-invasive assessment of tissue iron overload. *Hematology Am Soc Hematol Educ Program*. 2009;215-21.
- Saiviroonporn P, Viprakasit V, Boonyasirinant T, Khuhanapant A, Wood JC, Krittayaphong R. Comparison of the region-based and pixel-wise methods for cardiac T2* analysis in 50 transfusion-dependent Thai thalassemia patients. *J Comput Assist Tomogr*. 2011 May-Jun;35(3):375-81.
- Anderson LJ, Holden S, Davis B, Prescott E, Charrier CC, Bunce NH, et al. Cardiovascular T2-star (T2*) magnetic resonance for the early diagnosis of myocardial iron overload. *Eur Heart J*. 2001 Dec;22(23):2171-9.
- Brewer CJ, Coates TD, Wood JC. Spleen R2 and R2* in iron-overloaded patients with sickle cell disease and thalassemia major. *J Magn Reson Imaging*. 2009 Feb;29(2):357-64.
- Argyropoulou MI, Astrakas L. MRI evaluation of tissue iron burden in patients with beta-thalassemia major. *Pediatr Radiol*. 2007 Dec;37(12):1191-200; quiz 308-9.
- Cohen AR, Galanello R, Pennell DJ, Cunningham MJ, Vichinsky E. Thalassemia. *Hematology Am Soc Hematol Educ Program*. 2004;14-34.
- Ghugre NR, Gonzalez-Gomez I, Butensky E, Noetzli L, Fischer R, Williams R, et al. Patterns of hepatic iron distribution in patients with chronically transfused thalassemia and sickle cell disease. *Am J Hematol*. 2009 Aug;84(8):480-3.
- Hankins JS, McCarville MB, Loeffler RB, Smeltzer MP, Onciu M, Hoffer FA, et al. R2* magnetic resonance imaging of the liver in patients with iron overload. *Blood*. 2009 May 14;113(20):4853-5.
- Alustiza JM, Castiella A, De Juan MD, Empanaza JI, Artetxe J, Uranga M. Iron overload in the liver diagnostic and quantification. *Eur J Radiol*. 2007 Mar;61(3):499-506.
- Taher A, El Rassi F, Isma'eel H, Koussa S, Inati A, Cappellini MD. Correlation of liver iron concentration determined by R2 magnetic resonance imaging with serum ferritin in patients with thalassemia intermedia. *Haematologica*. 2008 Oct;93(10):1584-6.
- Noetzli LJ, Carson SM, Nord AS, Coates TD, Wood JC. Longitudinal analysis of heart and liver iron in thalassemia major. *Blood*. 2008 Oct 1;112(7):2973-8.
- Kontoghiorghe GJ. A new era in iron chelation therapy: the design of optimal, individually adjusted iron chelation therapies for the complete removal of iron overload in thalassemia and other chronically transfused patients. *Hemoglobin*. 2009;33(5):332-8.
- Berdoukas V, Choularas G, Moraitis P, Zannikos K, Berdoussi E, Ladis V. The efficacy of iron chelator regimens in reducing cardiac and hepatic iron in patients with thalassemia major: a clinical observational study. *J Cardiovasc Magn Reson*. 2009;11(1):20.
- Carpenter JP, Pennell DJ. Role of T2* magnetic resonance in monitoring iron chelation therapy. *Acta Haematol*. 2009;122(2-3):146-54.
- Westwood M, Anderson LJ, Firmin DN, Gatehouse PD, Charrier CC, Wonke B, et al. A single breath-hold multiecho T2* cardiovascular magnetic resonance technique for diagnosis of myocardial iron overload. *J Magn Reson Imaging*. 2003 Jul;18(1):33-9.
- Maris TG, Papakonstantinou O, Chatzimanoli V, Papadakis A, Pagonidis K, Papanikolaou N, et al. Myocardial and liver iron status using a fast T2* quantitative MRI (T2*QMRI) technique. *Magn Reson Med*. 2007 Apr;57(4):742-53.
- Pepe A, Positano V, Santarelli MF, Sorrentino F, Cracolici E, De Marchi D, et al. Multislice multiecho T2* cardiovascular magnetic resonance for detection of the heterogeneous distribution of myocardial iron overload. *J Magn Reson Imaging*. 2006 May;23(5):662-8.
- McCarville MB, Hillenbrand CM, Loeffler RB, Smeltzer MP, Song R, Li CS, et al. Comparison of whole liver and small region-of-interest measurements of MRI liver R2* in children with iron overload. *Pediatr Radiol*. Aug;40(8):1360-7.
- Positano V, Salani B, Pepe A, Santarelli MF, De Marchi D, Ramazzotti A, et al. Improved T2* assessment in liver iron overload by magnetic resonance imaging. *Magn Reson Imaging*. 2009 Feb;27(2):188-97.
- Porter JB, Shah FT. Iron overload in thalassemia and related conditions: therapeutic goals and assessment of response to chelation therapies. *Hematol Oncol Clin North Am*. 2010 Dec;24(6):1109-30.
- Viprakasit V, Lee-Lee C, Chong QT, Lin KH, Khuhanapant A. Iron chelation therapy in the management of thalassemia: the Asian perspectives. *Int J Hematol*. 2009 Nov;90(4):435-45.
- Positano V, Salani B, Scattini B, Santarelli MF, Ramazzotti A, Pepe A, et al. A robust method for assessment of iron overload in liver by magnetic resonance imaging. *Conf Proc IEEE Eng Med Biol Soc*. 2007;2007:2895-8.
- Kirk P, He T, Anderson LJ, Roughton M, Tanner MA, Lam WW, et al. International reproducibility of single breathhold T2* MR for cardiac and liver iron assessment among five thalassemia centers. *J Magn Reson Imaging*. Aug;32(2):315-9.
- St Pierre TG, Clark PR, Chua-Anusorn W. Measurement and mapping of liver iron concentrations using magnetic resonance imaging. *Ann N Y Acad Sci*. 2005;1054:379-85.